Pick up the Warm Up

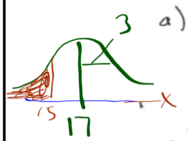
(1) The weights of suitcases at an airport are normally distributed with a mean of 17 kg and $\sigma = 3$ kg

a) How many of the 300 suitcases per hour would you expect to be lighter than 15 kg?

b) 4% of the suitcases on any day are rejected because they exceed the weight limit. What is the weight limit?

(1) The weights of suitcases at an airport are normally distributed with a mean of 17 kg and $\sigma = 3$ kg

a) How many of the 300 suitcases per hour would you expect to be lighter than 15 kg?



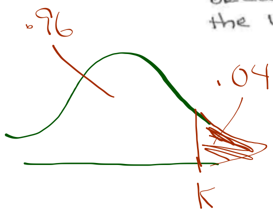
$P(x<15) = 0.252...$

↳ normalcdf $(-1000, 15, 17, 3)$

$(.252...)(300)$

75.7 suitcases

b) 4% of the suitcases on any day are rejected because they exceed the weight limit. What is the weight limit?

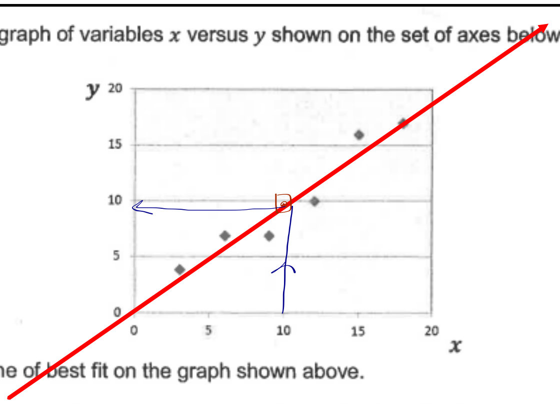


$P(x<K) = .96$

$K = 22.3$ kg

↳ invNorm$(.96, 17, 3)$

(2) Consider the graph of variables x versus y shown on the set of axes below:



a. Draw a line of best fit on the graph shown above.

b. Circle the correlation coefficient shown below that best illustrates the relationship shown between the two sets of data x and y.

$r \approx 0$ $r \approx -0.96$ $r \approx 0.96$ $r \approx 0.24$

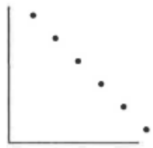
c. Use the line of best fit drawn in part (a) above to estimate a value of y corresponding to an x value of 10.

≈ 9

(3)

Match the letter of the appropriate correlation coefficient with the graphs shown below:

Graph 1:



$r \approx -1.0$

Graph 2:

$r \approx 0$

Graph 3:

0.7

A. $r \approx 0$ B. $r \approx +1.0$ C. $r \approx -1.0$ D. $r \approx +0.7$ E. $r \approx -0.7$

(4)

Ten middle years students were measured for height (h) and arm span (a). The results are shown in the table below:

Height: h (cm)	Arm Span: a (cm)
152	154
156	154
160	158
164	166
166	163
166	167
170	172
175	174
177	178
180	178

166.6
167

a. Calculate $\bar{h}$ and $\bar{a}$. $\bar{h} = 167$ cm $\bar{a} = 166$ cm

b. Determine the correlation coefficient between h and a. $r = 0.98$

c. Use words to describe the relationship between h and a.

There is a strong, positive, correlation between Height and arm span for this group.

As the height increases, the arm span increase.

Slope-Intercept

$y = mx + b$

$(18, -6)$ $(9, 1)$

$m = \frac{-6 - 1}{18 - 9}$

$= \frac{-7}{9}$

Point - Slope

$y - y_1 = m(x - x_1)$

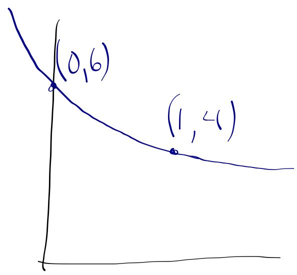
$y - 1 = \frac{-7}{9}(x - 9)$ ✓

$y + 6 = \frac{-7}{9}(x - 18)$ ✓

Go over

HW

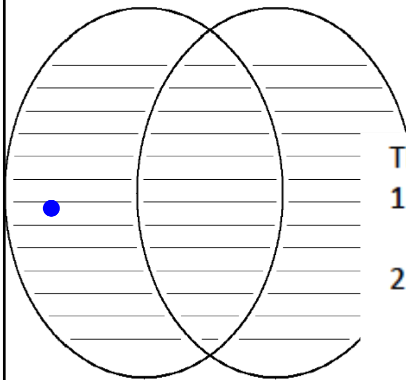
#4 $f(x) = p(0.5)^x + q$



$6 = p(0.5)^0 + q$ $\quad$ $4 = p(0.5)^1 + q$

$6 = p + q$ $\quad$ $4 = .5p + q$

Association vs. Causation



Three reasons why two variables may be associated:

1. An experiment proves that one variable causes the other. This is a true causation.
2. A third variable impacts both variables making it appear that there is a cause and effect relationship. This is an association.
3. There is a correlation but it is just coincidental due to a very small sample size. This is an association.

AIM

Calculate the correlation coefficient, "by hand"using the formula itself.

There are a few methods to calculate the correlation coefficient, r. The one we will be looking at was invented by someone called Pearson, and its full title is......

Pearson's Product Moment Correlation Coefficient

will also be in the Ch 11 packet

Two Variable Statistics – Day 2
Class Notes

Calculate the Linear Correlation Coefficient by Hand

Terminology → or explanatory

$\bar{x}$ mean of the independent variable

$\bar{y}$ mean of the dependent variable
↳ or response

$(x_i - \bar{x})$ the deviation from the mean of the independent variable.

$(y_i - \bar{y})$ the deviation from the mean of the depend. variable.

$(x_i - \bar{x})^2$ the Square of the deviation from the mean of the indep. variable.

$\Sigma(x_i - \bar{x})^2$ the Sum of the squares of the deviation of the independent vari

of the deviation

$\Sigma(y_i - \bar{y})^2$ sum of squares ~~of deviation of the~~

of the depend. variable.

$(x_i - \bar{x})(y_i - \bar{y})$ the Product of the deviations from the means of both variables

$\Sigma(x_i - \bar{x})(y_i - \bar{y})$ the sum of the products of deviations of both means,

$$r = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\Sigma(x_i-\bar{x})^2 \times \Sigma(y_i-\bar{y})^2}} = \frac{\text{critical total \#1}}{\sqrt{\text{crit total \#2} \times \text{crit total \#3}}}$$

Example

Calculate the correlation coefficient
Showing all critical totals

Distance from the statue	Price of the Bottle
10 metres	$2.80
50 metres	$2.70
80 metres	$2.60
100 metres	$2.40
130 metres	$2.20
170 metres	$2.00

$\bar{x} = 90$

$\bar{y} = 2.45$

Example

Calculate the correlation coefficient
Showing all critical totals

Distance from the statue	Price of the Bottle
10 metres	$2.80
50 metres	$2.70
80 metres	$2.60
100 metres	$2.40
130 metres	$2.20
170 metres	$2.00

Needs:

$\bar{x}$

$\bar{y}$

Example

Calculate the correlation coefficient
Showing all critical totals

Distance from the statue	Price of the Bottle
10 metres	$2.80
50 metres	$2.70
80 metres	$2.60
100 metres	$2.40
130 metres	$2.20
170 metres	$2.00

Needs

$\bar{x} = 90$ meters

$\bar{y} = \$2.45$

→ GDC
to get one critical
total at a time

$r = \frac{\quad}{\sqrt{\quad}}$

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{}}$$

$$\bar{x} = 90$$
$$\bar{y} = 2.45$$

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \times \sum (y_i - \bar{y})^2}} = \frac{-86}{\sqrt{16200 \cdot .475}}$$

$$= -.980$$

$$= \frac{}{\sqrt{}}$$

↖ leave blank for now

$$r = \frac{-86}{(16200)(0.475)}$$

$$r = -0.980$$

For IB exams:

a) On the IB exam, you would only use your calculator to quickly calculate **r**

b) If you use correlation on your project, you would have to include a calculation by hand
(with the help of a spreadsheet most likely.
(checked by a calculator, perhaps)

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \times \sum (y_i - \bar{y})^2}} =$$

$\bar{x} = 90$

$\bar{y} = 2.45$

$$= \frac{(10-90)(2.80-2.45) + (50-90)(2.70-2.45) + \ldots}{\sqrt{\left[(10-90)^2 + (50-90)^2 + \ldots\right]\left[(2.80-2.45)^2 + (2.70-2.45)^2 + \ldots\right]}}$$

BB

Assignment Day #2 is a worksheet

Due tomorrow.

Optional Extra Practice Problems for tomorrow's 15 to 20 minute quiz on Normal Distribution

Answers are posted along with the others. These are not required to be turned in.

p. 312 Review Set A....1, 3, 6 and Set B... 2, 5